

MASSACHUSETTS INSTITUTE OF TECHNOLOGY  
ARTIFICIAL INTELLIGENCE LABORATORY

A. I. Memo No. 491

September 1978

**BANDPASS CHANNELS, ZERO-CROSSINGS, AND EARLY  
VISUAL INFORMATION PROCESSING**

by

**D. Marr, T. Poggio and S. Ullman**

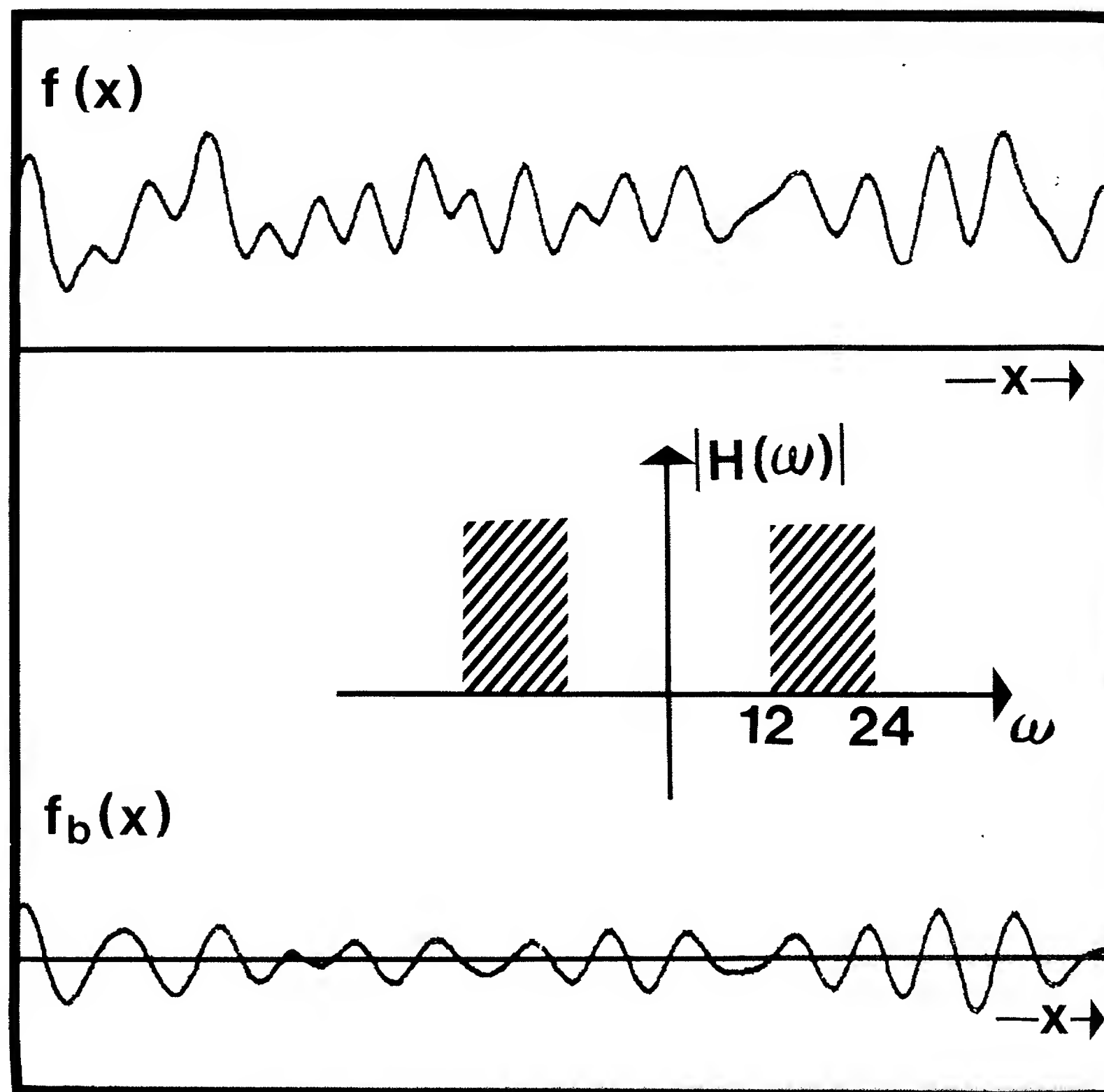
**ABSTRACT:** A recent advance by B.F. Logan in the theory of one octave bandpass signals may throw new light on spatial-frequency-tuned channels in early visual information processing.

This work was conducted at the Max-Planck-Institut fuer Biologische Kybernetik in Tuebingen, and at the Artificial Intelligence Laboratory, a Massachusetts Institute of Technology research program supported in part by the Advanced Research Projects Agency of the Department of Defense, and monitored by the Office of Naval Research under contract number N00014-75-C-0643. D.M. and S.U. were also supported by NSF contract number 77-07569-MCS.

In most approaches to computer vision, an important preliminary computation is the localization of discontinuities in image intensity. This can be achieved by finding peaks in the first directional derivative of intensity, or equivalently, zero-crossings in the second directional derivative. The latter quantity may be obtained by convolving the image with a bar-shaped mask, which approximates the second directional derivative at its particular scale. By using a range of mask sizes, one can begin to deal with the wide range of scales over which changes take place in a natural image (see Marr, 1976, p. 488).

These ideas begin to account, on purely information processing grounds, for the presence of frequency-tuned channels in early human vision (Campbell & Robson, 1968). Recent work by Wilson & Giese (1977) shows that such channels can be realised by linear units with bar-shaped receptive fields, reminiscent of the simple cells that Hubel & Wiesel (1962) have described. Marr & Poggio's (1977, 1979) recent theory of stereopsis is, for example, conceived within this framework, and assumes that the elements that are matched between the two images are equivalent to the zero-crossings in bar-mask outputs. The object of this note is to point out that very recent advances in information theory provide fascinating additional theoretical support for this framework.

The advance in question is a theorem due to Logan (1977), who showed that if a one-dimensional analytic function is (a) bandpass of bandwidth one octave or less, and (b) has no *free zeroes*, i.e. complex



1. The meaning of Logan's (1977) theorem. (a) shows a stochastic gaussian signal  $f(x)$ , band-limited by  $\omega = 24$ , and (c) exhibits the result  $f_b(x)$  of filtering (a) through an ideal one-octave bandpass filter. The modulus of its transfer function is shown in (b). Since (c) has a bandwidth of one octave, and it has no zeros in common with its Hilbert transform, Logan's theorem tells us that (c) is determined, up to a multiplicative constant, by its zero-crossings alone. The aspect of Logan's result that is important for this article is that under the right conditions, zero-crossings alone are very rich in information.

zeroes in common with its Hilbert transform, then the function is completely determined (up to an overall multiplicative constant) by its (real) zero-crossings (see figure 1). Condition (a) is critical, but condition (b) can for practical purposes be ignored, since it is almost always satisfied except by pathological signals.

If one translates this result into the context of early visual processing, its meaning is this. We have already seen that the basic idea, of using zero-crossings in bar-mask convolutions from which to generate a primitive description of the image, has a strong physical motivation. Logan's result tells us that, if the bar-mask operators are band-pass with a bandwidth of not more than one octave, then the zero-crossings alone are so rich in information that they determine essentially completely the convolution values (taken along a scan-line perpendicular to the mask's orientation).

Another basic question which Logan's result may illuminate is, why should the channels used in early visual processing be orientation-dependent? Why not compute one's primitive description directly from circularly symmetric masks, like the receptive fields of retinal ganglion cells? Imagine that one wishes to reconstruct a two-dimensional array from the zero-crossings along a family of scan-lines that cover the plane. Logan's result tells us that this is in general impossible from the zero-crossings alone unless the array values along each scan-line are bandpass with bandwidth less than an octave. It is not enough that the two-dimensional array be bandpass in two dimensions with bandwidth less than an octave (as a ring in the  $(\omega_x, \omega_y)$  plane of

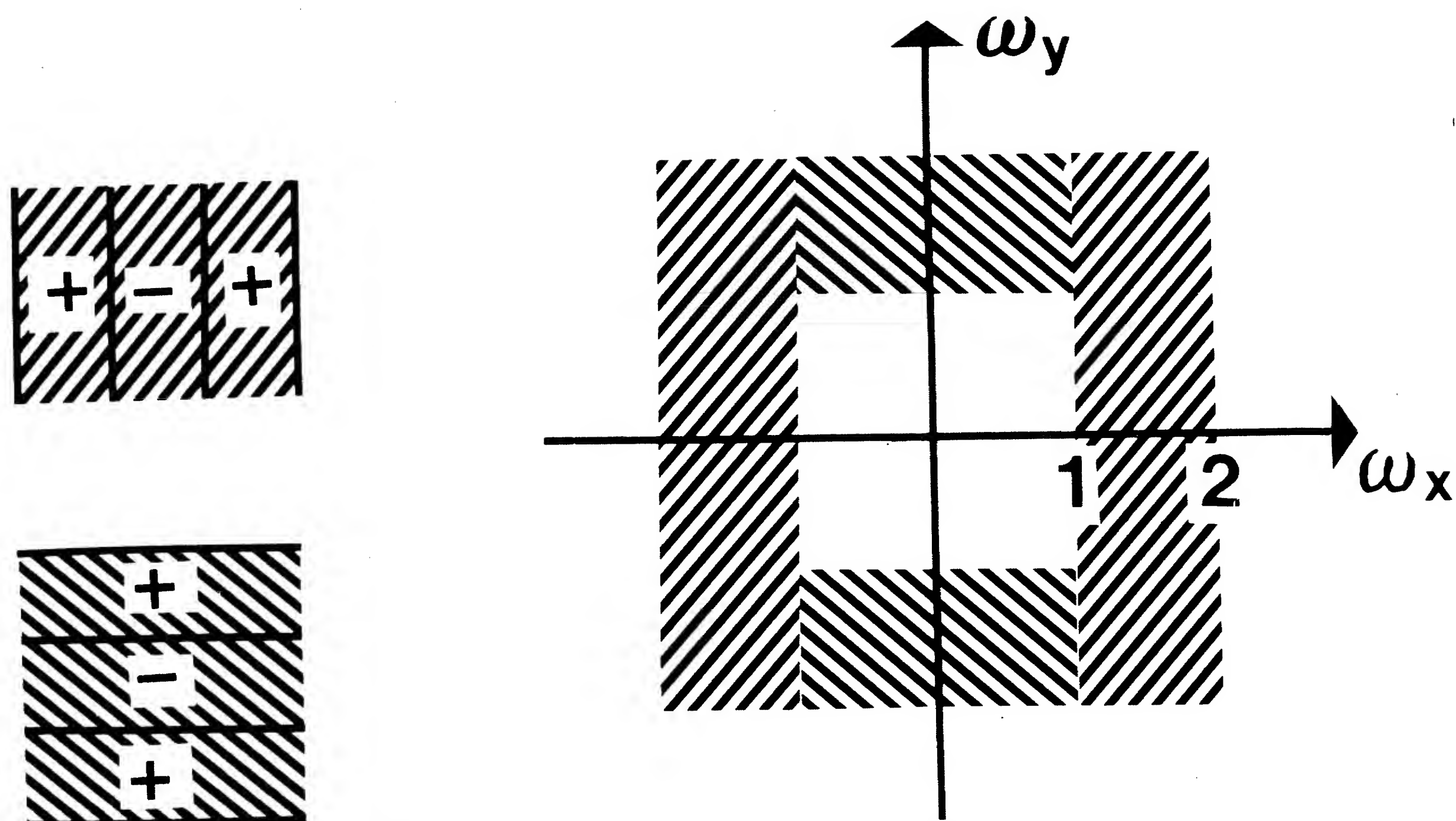


width one octave). An image filtered through a (bandpass) bar-shaped mask is bandpass on each scan-line perpendicular to the mask's orientation; an image filtered through a (bandpass) circularly-symmetric mask is band-limited but not bandpass along any scan-line. This follows from the fact that the Fourier transform along (for instance) the  $x$ -axis of an image filtered through a bandpass "ring" is essentially the projection of the two-dimensional Fourier transform on  $\omega_x$ , and is therefore not bandpass. We may conclude that a commitment to one-dimensional techniques, (i.e., zero-crossings along scan-lines), obliges one to use orientation-dependent masks.

This argument, however, gives us no clue about the number of orientations that one should use. For reconstructing the image, the Logan approach provides a lower bound of two orientations, together with an adequate set of mask sizes (see figure 2).

In its extreme form, our thesis may be summarized as follows. In order to construct a faithful representation of the image using only zero-crossings, it is necessary to filter it through a set of independent bandpass channels with one octave bandwidth. Hence the masks (or receptive fields) that approximate the second directional derivative operator should, as closely as possible, be bandpass with one octave bandwidth. Such a system would allow the recovery of sharp intensity changes directly from the mask outputs, while providing the necessary basis for the recovery of arbitrary intensity profiles.

What experimental evidence is there that our thesis is relevant to biological visual systems? As we mentioned earlier, Logan's free zero



2. On the left are shown bar-shaped masks at the vertical and horizontal orientations, and on the right, the amplitude of their (idealized) transfer functions. The bandwidth shown here is one octave, the maximum value for which Logan's theorem applies. (In practise, an ideal one-octave bandwidth requires side-lobes in the "receptive field".) If for each mask, zero-crossings are found along scan-lines lying perpendicular to the mask's orientation, these zero-crossings contain full information about that part of the image whose spectrum falls within the shaded region (on the right) of the Fourier plane. The remaining regions of the Fourier plane can be covered by similar masks of different sizes.

Interestingly, if one uses masks constructed from the difference of two gaussian curves, their Fourier transforms behave like  $\omega^2$ , for values of  $\omega$  that are small compared with  $\sigma$ . In other words, they approximate a second derivative operator.

condition will almost always be satisfied in practice. The critical condition concerns the bandwidth. There is ample evidence for the existence in the human visual system of independent, spatial-frequency-tuned bandpass channels, of about one-octave bandwidth. Precise estimates of the bandwidth vary considerably, however, ranging from very narrow (0.5 octaves, Sachs, Nachmias & Robson 1971) to very large (Kulikowski & King-Smith 1973; Shapley & Tolhurst 1973) values. More recent approaches based on spatial probability summation allow most of the existing psychophysical data to be fitted using medium bandwidth channels. Graham's (1977, Figure 4) estimate of channel bandwidth half-peak sensitivity is about 0.5 octaves, whereas the especially convincing estimates of Wilson & Giese (1977) hover around an octave and a half (see also Legge 1978). In any case, the channels are not the ideal one-octave bandpass filters that Logan's theorem requires. There is unfortunately little available information about channel characteristics in their normal (suprathreshold) conditions, although there are hints that their bandwidth may then be somewhat narrower (Cowan 1977, Figure A12). Furthermore, it seems likely that Logan's one-octave condition may be relaxed. (The average failure rate at 1.5 octaves is probably around 8%). In any case, it becomes of considerable interest to determine the channel bandwidths under suprathreshold conditions.



## References

- Campbell, F.W. and Robson, J.G. 1968. Applications of Fourier analysis to the visibility of gratings. *J. Physiol. (Lond.)* 197, 551-556.
- Cowan, J.D. 1977. Some remarks on channel bandwidths for visual contrast detection. Figure A12. *Neurosciences Res. Prog. Bull.* 15, 492-517.
- Graham, N. 1977. Visual detection of aperiodic spatial stimuli by probability summation among narrowband channels. Figure 4. *Vision. Res.* 17, 637-652.
- Hubel, D.H. & Wiesel, T.N. 1962. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol. (Lond.)* 160, 106-154.
- Kulikowski, J.J. & King-Smith, P.E. 1973. Spatial arrangement of line, edge, and grating detectors revealed by subthreshold summation. *Vision. Res.* 13, 1455-1478.
- Legge, G.E. 1978. Sustained and transient mechanisms in human vision: temporal and spatial properties. *Vision. Res.* 18, 69-81.
- Logan, B.F., Jr. 1977. Information in the zero-crossings of bandpass signals. *Bell System Technical Journal* 56, 487-510.
- Marr, D. 1976. Early processing of visual information. *Phil. Trans. R. Soc. B (Lond.)* 275, 483-524.
- Marr, D. & Poggio, T. 1977. Theory of human stereopsis. *J. opt. Soc. Am.* 67, 1400.
- Marr, D. & Poggio, T. 1979. A computational theory of human stereo vision. Submitted to the Royal Society for publication.
- Pollen, D.A., Lee, J.R. & Taylor, J.H. 1971. How does the striate cortex begin the reconstruction of the visual world? *Science* 173, 74-77.
- Sachs, M.B., Nachmias, J. & Robson, J.G. 1971. Spatial-frequency channels in human vision. *J. opt. Soc. Amer.* 61, 1176-1186.
- Shapley, R.M. & Tolhurst, D.J. 1973. Edge detectors in human vision. *J. Physiol. (Lond.)* 229, 165-183.
- Wilson, H.R. & Giese, S.C. 1977. Threshold visibility of frequency gradient patterns. *Vision. Res.* 17, 1177-1190.